Urban Mobility Flows from Mobile Phone Data

Diala Naboulsi, Marco Fiore, Razvan Stanica

Urban Mobility Flows from Mobile Phone Data

DialaNaboulsi[1], Marco Fiore [2], RazvanStanica[1]
[1] Université de Lyon, INSA Lyon, CITI-INRIA UrbaNet, Villeurbanne, France
[2] CNR – IEIIT, Torino, Italy
Contact: diala.naboulsi@insa-lyon.fr

Context and Motivation

Understanding human movements is critical for different scientific domains. In orderto deploy efficient networking solutions, a clear view of human mobility patterns is required. The same applies for urban planning, where the global mobility flows can determine the optimal deployment of infrastructure. Human mobility also plays a major role when analyzing the ways diseases can spread in a population.

Significant research efforts have been conducted in this direction, aiming at understanding how people move as a first step, and proposing models of such mobility as a second step. Recently, as people are more and more connected, network traces have received particular attention as a source of information about human mobility at large scales.

However, previous studies have focused on developed countries, and whether the observed patterns and models are applicable to developing countries remains an open question, due to differences in the lifestyle, country's infrastructure and modes of transportation. Indeed, a clear understanding of human movements would be crucial for the progress of such countries, especially in highly populated urban regions where new transportation infrastructures are being deployed.

In this work, we explore Call Detail Records (CDR) of Orange customers in Abidjan, the economic capital of Ivory Coast. The dataset, made available within the context of the D4D Challenge, provides the position of each caller – approximated as the base station's location – at every time he/she initiates a call or sends an SMS. We start by analyzing the temporal, spatial and geographical characteristics of the calls, which allows us to capture differences between distinct times of the day and different days of the week over multiple geographical regions of the city. We propose a method to distinguish between typical and outlying behaviors in the CDR dataset, enabling the detection of special events such as the New Year's Eve and football games played during the Africa Cup of Nations. Our approach also allows us to infer which moments can be aggregated in order to characterize macroscopic mobility flows that provide a view of the globaland local mobility flows in Abidjan, as well as of their daily evolution.

Methodology

In our analysis we study datasets provided by Orange within the context of the D4D Challenge, based on the Call Detail Records of 5 million anonymized Orange customers in Ivory Coast. The information obtained spans over 5 months, from December 5th, 2011 until April 22nd, 2012. We focus on two of the four datasets provided, detailed below.

Dataset D1: Antenna-to-antenna traffic. This dataset includes the call traffic volume exchanged between any two base stations in Ivory Coast on an hourly basis. It

provides both the number of calls and their total duration for the whole observation period.

Dataset D2: Individual trajectories. For each two weeks of the observation period, this dataset provides the CDR of 50,000 individuals randomly chosen over the whole Ivory Coast Orange customer population.

Our study focuses on the city of Abidjan, the economic capital of Ivory Coast. Thus, we filter both D1 and D2 by keeping only the information involving the antennas in Abidjan. This leaves us with information about 364 antennas out of the 1231 antennas covering the whole country.

We remark that, in the following, we will use the term *snapshot* to refer to data aggregated over each one-hour interval, and the term *call* to denote a call or an SMS indifferently.

Results

From the two datasets considered, D1 cannot be used as the main source for an O/D matrix. The reason is that, although it provides us with a view of the entire traffic over the cellular network, the dataset D1 is focused on values aggregated per base station, without any information on the individual trajectories. Therefore, in order to construct the mobility flows in the Abidjan area, we make use of the dataset D2 that contains individual information for a subset of the users.

However, using only a subset of the individuals, sampled by the operator following an unknown distribution, can reduce the accuracy of the mobility matrices. To alleviate this problem, we decided to aggregate multiple days, which allows us to follow a larger set of individuals; for example we aggregate multiple Tuesdays at time 10:00, therefore creating a *classical 10am Tuesday*. Nevertheless, continuing on this example, some of the Tuesdays available in the dataset might present a non-typical behavior during a certain time interval. Distinguishing these situations is important if we want to build typical O/D matrices. In order to establish which days can be aggregated, we use the dataset D1, which contains the complete data.

We define two metrics, which allow us to consider variations both in call volume and user distribution. The second metric is essential in this analysis, as it helps us detect moments when the population is distributed in an atypical manner. This technique allows us to detect special events such as the New Year's Eve, but also 5 out of 6 football games of the Africa Cup of Nations 2012, hosted by Gabon and Equatorial Guinea, in which the Ivory Coast participated.

Once these *outlying* behaviors are filtered, we can build the O/D matrix using dataset D2. Some of the observations made on these matrices are summarized below for a classical Thursday:

- as early as 6am, minor movements can be distinguished going towards the city center;
- at 8am we distinguish a major mobility trend, oriented towards the working and studying areas of the city;
- from 1pm, mobility flows begin leaving the city center, the main working area, probably as a consequence of the important mass of part-time workers that exist in Abidjan;
- the movement towards the residential areas continues through the afternoon, with a peak around 7pm.

- finally, the city calms down and mobility highly reduces after 10pm.

Conclusion

In this paper, we analyze Call Detail Records of mobile phone users in the city of Abidjan. We introduce a method that allows distinguishing between typical and special calling behaviors of the population. We extract the global mobility flows across the whole city, which we prove to reflect the dynamics of the lifestyle in Abidjan. We believe that our results can help solve important problems in the city. They can be exploited to improve the public transportation services, adapting the paths taken by buses and taxis as well as their number to the mobility flows. They can also be used when considering traffic problems and the road infrastructure, showing where new roads are needed: our results clearly confirm the need for the construction of the third bridge linking the northern and southern parts of the city. Mobile phone services can be ameliorated as well based on our results, by adapting them to the macroscopic movements detected taking into account the geographical locations where people mostly cluster at different times of the day. As for the accuracy of the obtained flows, we remark that it can be improved by considering larger, more complete datasets. Finally, our results are limited to one city in Ivory Coast, a similar study can be achieved for the other areas of the country, also it can be interesting to check the mobility flows between different regions of Ivory Coast when considering the improvement of services in the country as a whole.